

Patterning the insect eye: from stochastic to deterministic mechanisms

Haleh Ebadi,¹ Michael Perry,² Keith Short,² Konstantin

Klemm,^{3,4} Claude Desplan,² Peter F. Stadler,¹ & Anita Mehta^{1,5}

¹Bioinformatics, Institute for Computer Science, Leipzig University, Härtelstrasse 16-18, 04107 Leipzig, Germany

²Department of Biology, New York University, 1009 Silver Center,
100 Washington Square East, New York, NY 10003, USA.

³Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Palma de Mallorca, E-07122, Spain

⁴Department of Computer Science, School of Science and Technology,
Nazarbayev University, Astana, Republic of Kazakhstan, 010000.

⁵Theoretical Sciences, S N Bose National Centre,
Block JD Sector III Salt Lake, Calcutta 700 106, India

ABSTRACT:

While most processes in biology are highly deterministic, stochastic mechanisms are sometimes used to increase cellular diversity, such as in the specification of sensory receptors. In the human and *Drosophila* eye, photoreceptors sensitive to various wavelengths of light are distributed randomly across the retina. Mechanisms that underlie stochastic cell fate specification have been analysed in detail in the *Drosophila* retina. In contrast, the retinas of another group of dipteran flies exhibit highly ordered patterns. Species in the Dolichopodidae, the “long-legged” flies, have regular alternating columns of two types of ommatidia (unit eyes), each producing corneal lenses of different colours. Individual flies sometimes exhibit perturbations of this orderly pattern, with “mistakes” producing changes in pattern that can propagate across the entire eye, suggesting that the underlying developmental mechanisms follow local, cellular-automaton-like rules. We hypothesize that the regulatory circuitry patterning the eye is largely conserved among flies such that the difference between the *Drosophila* and Dolichopodidae eyes should be explicable in terms of relative interaction strengths, rather than requiring a rewiring of the regulatory network. We present a simple stochastic model which, among its other predictions, is capable of explaining both the random *Drosophila* eye and the ordered, striped pattern of Dolichopodidae.

INTRODUCTION

The development of multicellular animals is highly reproducible, with deterministic and orderly processes generating reliable outcomes. Segment boundaries form in the proper place and cell types are set aside in specific proportions in differentiating tissues. Underlying these seemingly precise developmental outcomes, though, are inherently stochastic transcriptional events, *e.g.* decisions to express or not express key regulators of cell fate (1, 2). Varying amounts of activating or repressive input can bias these decisions strongly one way or the other, producing seemingly deterministic on or off outcomes, resulting in distinct boundaries and specific spatial patterns (3, 4). The distribution of these inputs depends largely on lineage and positional information within an embryo or tissue. In another class of cell fate decisions, stochastic cell-intrinsic mechanisms instead produce particular probabilities of taking one fate or another in otherwise equivalent cells (5). In their own way, these stochastic decisions are highly regulated to take place in specific tissue types and to produce reliable proportions of one cell fate vs. another.

How such probabilistic patterning mechanisms might be switched between stochastic and deterministic is a question to which the tools of statistical physics can meaningfully be applied. An example of stochastic patterning occurs in the fly eye (5, 6), a complex organ whose development has been the subject of great scrutiny (7, 8). Our interest focuses on the patterning of two photoreceptors (PRs) that are involved in colour vision; the two “inner PRs” R7 and R8 are randomly distributed across the retina (9), as seen via staining with antibodies against the green-sensitive or blue-sensitive photopigments, the Rhodopsins in R8 cells (Figure 1a). A similar stochastic pattern exists in R7 cells for two UV Rhodopsins, Rh3 and Rh4 (10). Stochastic on or off expression of the transcription factor Spineless in the R7 PR controls ommatidial type, and therefore the overall random pattern (5). In contrast, another group of flies in the family Dolichopodidae (referred to here as “Doli”) have ordered retinal patterning with alternating columns of ommatidia (the individual units of the adult compound eye) which produce two distinct corneal lens colours (Figure 1b). The patterning mechanisms that underlie both differentiation of PR types (*e.g.* R7 vs. R8) and stochastic patterning

across ommatidia have been shown to be largely conserved between *Drosophila* and butterflies (11, 12). Considering the apparent similarities between the *Drosophila* and Doli eye, it is tempting to suggest that the cell fate decisions involved in stochastic vs. non-stochastic patterning share the same underlying regulatory mechanisms with similar downstream effectors, but differ in a few critical upstream steps, and it might be possible to establish the rules that governed evolution from one mode of patterning to the other.

In this work, we present a simple mathematical model for such a regulatory mechanism, and compare our results with experimental data from the two fly species. Our model is also predictive and applicable to patterns observed in the eyes of other flies; we present predictions for the eyes of another species of Dolichopodidae that displays intermediate patterns as an example.

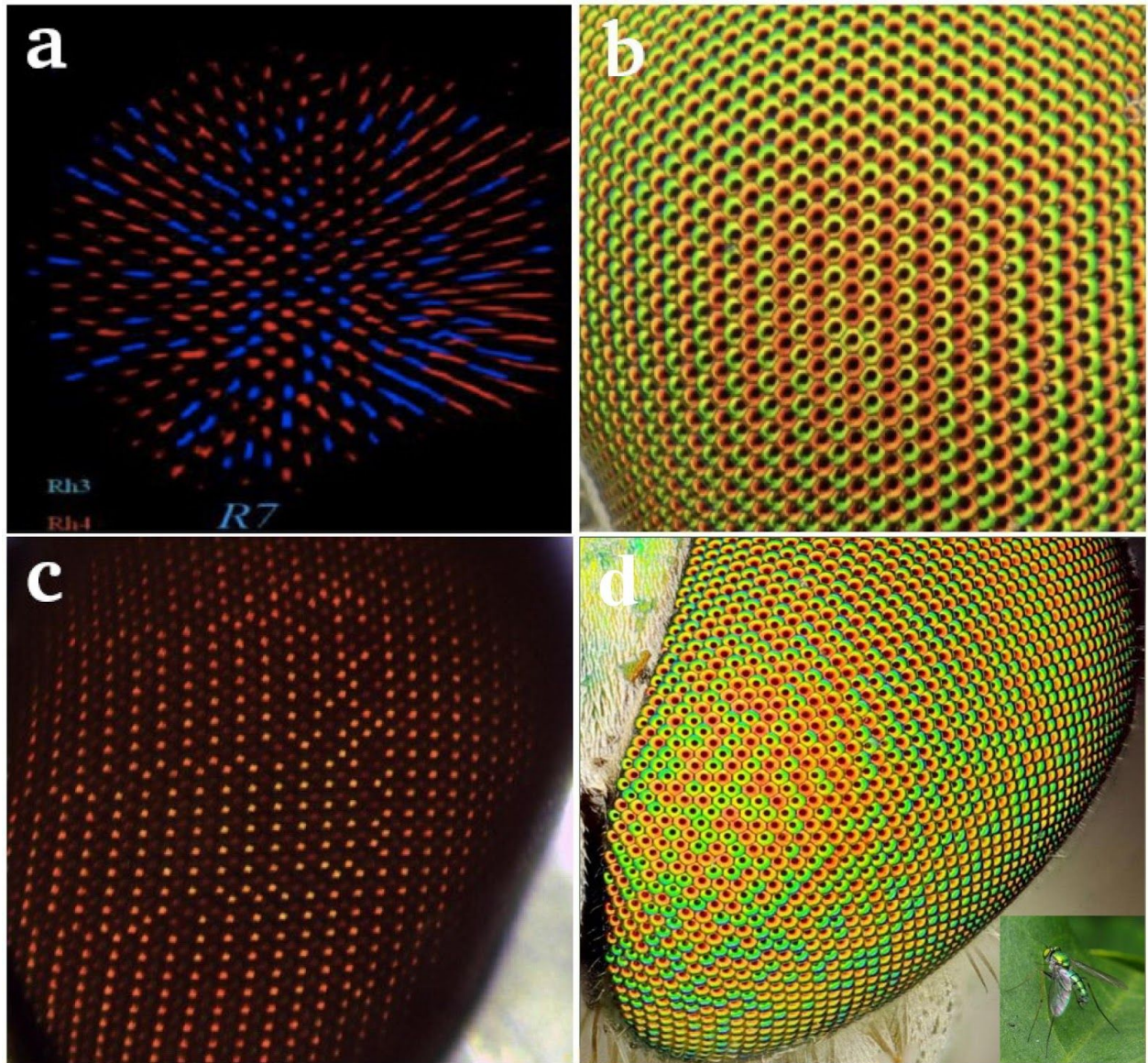


Figure 1: (a) Stochastic distribution of UV-sensitive Rhodopsin proteins (Rh3, blue and Rh4, red), in R7 cells of the *Drosophila* retina. (b) Striped pattern of red and yellow/green ommatidia in the *Dolichopodidae* eye. Note that isolated errors do not propagate. (c) *Doli* eye with several mistakes. Note that multiple errors lead to a switch of fate within a row, from right to left (posterior to anterior). (d) Eyes of a *Chrysosoma* sp., family Dolichopodidae, from Hawaii. Patterning is intermediate between ordered and stochastic.

The adult eye is a geometrically regular structure composed of hexagonal unit eyes packed into a grid. Patterning begins with the progression of the morphogenetic furrow, a posterior-to-anterior wave of differentiation. Sequential rounds of signalling produce ~25 highly ordered rows of ~ 30 ommatidia each to make up a total of 800 ommatidia per eye (13). Each ommatidium contains eight PRs and accessory cells: the six “outer PRs” (R1-R6) express a broad-spectrum Rhodopsin, Rh1, and are required for motion and dim light vision. The two “inner PRs” (R7 and R8) each express different Rhodopsins and are used for colour discrimination and polarized light vision (8, 14–16). A detailed mathematical model for much of the process of eye formation has recently been formulated (17). However, that model does not address the stochastic distribution of colour photoreceptors, which is the subject of this paper.

There are three main ommatidial subtypes in *Drosophila*, which are defined by the combination of Rhodopsin photopigments expressed in their R7 and R8 photoreceptors. Two of these, the so-called ‘pale’ and ‘yellow’ ommatidia, are randomly distributed across the retina in the ratio 35:65 (5, 9). The pale ommatidia express UV-sensitive Rh3 in R7 and blue-sensitive Rh5 in R8 and are used for the discrimination of short-wavelength light (5, 8). The yellow ommatidia express longer UV-sensitive Rh4 in R7 and green-sensitive Rh6 in R8 and are used for the discrimination of longer wavelengths (18) (Figure 1a). A third subtype found in the dorsal rim area (DRA) is used for the detection of the vector of light polarization (19). The stochastic distribution of yellow and pale ommatidia in *Drosophila* is controlled at a single upstream node in the retinal regulatory network by the stochastic expression of the transcription factor Spineless (Ss) in a subset of R7 cells (5). In *Doli*, where the patterning is highly ordered, Ss might also be responsible for Rhodopsin expression as the eyes appear to develop in highly similar ways; it also seems likely that many of the interactions in the eye regulatory network are conserved between *Doli* and *Drosophila* (11, 12). The generation of the very different patterns observed might thus be due to changes in the initial expression of Ss. In this paper, we present a simple mathematical model that captures the essence of these ideas, by attributing the diverse patterning in the two fly species to a single switching mechanism.

Model of retinal patterning.

Initial eye development proceeds via a complex set of interactions between cell signalling and changes in target gene expression as new cell types are sequentially recruited (13). After the progression of the morphogenetic furrow and recruitment of all cell types that will make up the adult retina, cell biological processes begin to shape and structure the ommatidia. The PRs produce microvillae that make up the rhabdomeres, the light gathering structures. The decision to express Ss (or not) determines the choice of Rhodopsins in the inner photoreceptors R7 and R8, and consequently, the colour sensitivity of the ommatidium. In *Drosophila*, this decision leads to a random distribution of Rhodopsins in pale or yellow ommatidia (5).

In contrast, *Doli* eyes instead show an orderly pattern of alternating columns, such as observed in the genus *Condylostylus* (Fig. 1b). Interestingly, we observed occasional perturbations in patterning in wild-collected *Doli* (Fig.1c). In some individuals, when multiple errors occur in adjacent or nearly adjacent ommatidia, errors in patterning sometimes propagate in an anterior direction from the initial column containing mistakes (see Fig. 1c for an example). In some animals many errors occurred in approximately the same column on the anterior-posterior axis in both retinas, suggesting a developmental cause such as thermal stress during a specific time during the migration of the morphogenetic furrow. Whatever the cause, the subsequent propagation of errors in the direction of the morphogenetic furrow suggests that initially local, cellular-automaton-like rules are at work.

With the idea that local signalling might set up more global patterns of alternating columns in mind, we made a key assumption: the regulatory circuitry being largely conserved among flies, any differences between fly species might be explicable in terms of relative strengths of interactions, rather than through an entire rewiring of the regulatory network itself. In this spirit, we assume the existence of a single gene product X that is required to activate the switch referred to above; thus, a (deterministic or stochastic) decision to express a specific Rhodopsin can be triggered only in the

presence of X . We also assume that the coupling strength of the factor X with the switching mechanism which eventually determines the eye colour, differs between fly species and can be influenced by changes in the unknown factor X and/or in the switching mechanism itself. Furthermore, the expression of X in an ommatidium should itself depend on the decisions made in the neighbouring ommatidium that developed just before it during the progression of the morphogenetic furrow; accordingly, in our model, we choose X at a given ommatidium to be the (weighted) average of its values in its preceding nearest neighbours. To sum up: X is a factor that diffuses with the furrow, whose values at a given ommatidium are correlated with those of its previous neighbours, and which is responsible for making local decisions at the point where photoreceptors acquire their identity.

These different values of X may, depending on their magnitudes, be able to activate the switch required to express a specific Rhodopsin. For the two cases of immediate interest, this switch is S_s whose expression levels determine whether ommatidia are yellow or pale in *Drosophila*, and perhaps whether stripes are green or red in *Doli*. However, for the general case, we define a conditional probability $P(S|X)$ which embodies the dependence of the expression of an arbitrary factor S in a given ommatidium on the expression level of the diffusing factor X introduced above. This includes the limiting case when (as in *Drosophila*), the colour decision becomes independent of X ; here, the equation $P(S|X) = P(S)$ applies.

Dynamically, pattern formation in the fly eye proceeds column by column: that is, spatial organisation and colour choices are made for every column of future ommatidia traversed by the morphogenetic furrow in its progression across the eye (8, 20). This implies a discrete temporal separation of the columns of ommatidia, which makes their colour decision the final step in their formation. The specific parameters clearly depend on the fly species. The *Drosophila* eye has random ordering with a bias for the yellow ommatidial type. *Dolis* instead have ordered retinal patterning, where columns of ommatidia with alternating corneal lens colours are found. This suggests that the colour

decision is made by a bistable switch.

Some clues about the nature of this switch can be found in Doli retinas that contain mistakes, where the regular ordering of red and green columns is sometimes interrupted. These errors can be propagated locally, sometimes over multiple subsequent columns, which suggests that the process that produces order is itself based on biasing an inherently stochastic decision toward a reliable outcome, but which can be interrupted by external conditions. For example, local propagation of errors makes it unlikely that the entire field is patterned using global retina-wide gradients of positional information. It is possible that perturbations during development might increase the frequency of mistakes during patterning near the morphogenetic furrow at the time of the perturbation, as suggested by the presence of multiple errors per column and in the same column in both eyes. These phenomena suggest that colour decisions in a column can have a local influence on their neighbours.

Taken together, these features suggest a unified theory that contains two ingredients: one, a stochastic choice element, and two, a correlation between the expression of colour choices in adjacent columns. In the case of *Drosophila*, the second of these can be set to zero to reflect the purely random spatial ordering of colours. Accordingly, in the following section, we first define our parameters in the context of the Doli eye and look at its dynamic formation, discussing the propagation of perturbations. Next, we describe the formation of the *Drosophila* eye in the context of our model.

Patterning of the Doli eye.

We assume that the choice of a colour in each column is essentially complete by the time the morphogenetic furrow moves on to the next column. Thus, we regard fate establishment as instantaneous on the timescale at which the full ordering process occurs in the next column. Consistent with this assumption, *Ss* expression in *Drosophila* starts almost immediately after the morphogenetic furrow (21). In Doli, we denote the two alternate colours, green and red, by 1 and 0 respectively. In our formalism, the colour choice of the ommatidium in the i^{th} row and j^{th} column of the embedding

hexagonal close-packed lattice is defined by the element a_{ij} (which is either 1 or 0) in an $n \times m$ matrix. Two competing effects determine the value of a given element: The first is a default probability of being in state 1 (green) for every element in a vertical column. The second is that of the subtype correlation between the ommatidium and its nearest neighbours in the preceding column.

Default probabilities.

In our model for Doli, the expression of Ss leads to ommatidia being green. We are thus led to define the default probability $p_j(S|X)$ of all sites in a column j to be green as:

$$p_j(S|X) = \begin{cases} P_0 & X_j \leq X_0 \\ 1 - P_0 & X_j > X_0 \end{cases}$$

(1)

where P_0 is a constant. For all values smaller than some threshold X_0 , the probability for the column to be green (*i.e.* for factor S to be expressed) is given by P_0 : when the threshold is exceeded, this changes to $1 - P_0$. The expression of S leads in its turn to a rapid change in the value of X over column $j+1$, which can be simply modelled by the following linear equation:

$$X_{j+1} = \gamma - \beta p_j(S|X) \quad (2)$$

where γ and β are constants. If β is positive, successive columns are anticorrelated, so that this models the case of Doli flies (Figure 1b), while a negative sign would be appropriate for a fly with a uniform retina. On the other hand, the constant γ is strictly positive.

In order to build up an extremely ordered alternating pattern such as that of Doli shown in Figure 1a, we can choose the constant P_0 to be very small; this guarantees almost

uniform colour in each column. For clarity, let us consider a sample choice of parameters to see how the equations evolve, viz. $P_0 = 0.0001$, $\beta = 8$, $\gamma = 10$, $X_0 = 5$. If, at the j^{th} column, X_j is greater than the threshold $X_0 = 5$, then Eq. 1 suggests that the default probability p_j of the column to be green is very close to 1. Eq. 2 then implies that the value of X at column $j + 1$, $X_{j+1} \sim 2$, which is less than the threshold X_0 . In turn, the first line of Eq. 1 yields $p_{j+1} = 0.0001$, i.e. column $j+1$ is red with very high probability. Ordered 'stripes' of red and green, as in Doli, are thus built across the eye as depicted in Figure 2a.

This formalism is, however, much more general than the simple modelling of ordered Doli. For instance, flies such as the intermediate stochastic/ordered fly in Figure 1d have patterns on a more uniform rather than an alternating background. Such a uniform background can be generated by a negative sign of β . In this case, if the first column has an above-threshold value of X , this value is only added to in successive columns, because of the negative sign of β in Eq. 2. That is, all successive columns will also have values of X above threshold, thereby generating a background of uniform, rather than alternating colour. Flies such as the *Chrysosoma* sp fly can then be modelled as perturbations of this model 'uniform' fly, as shown in Figure 2d.

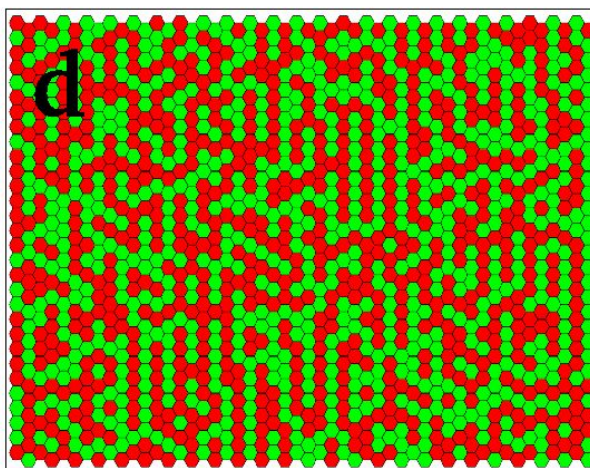
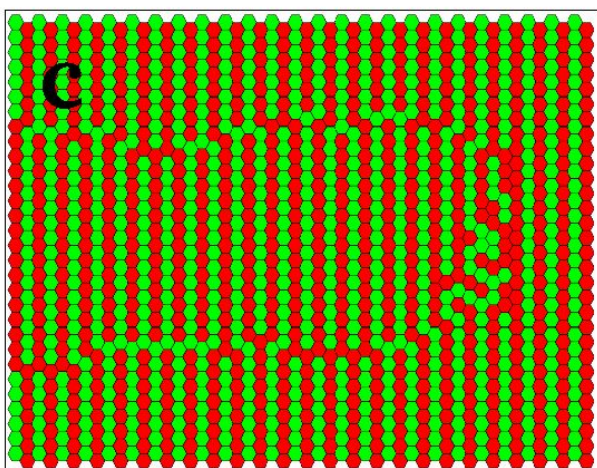
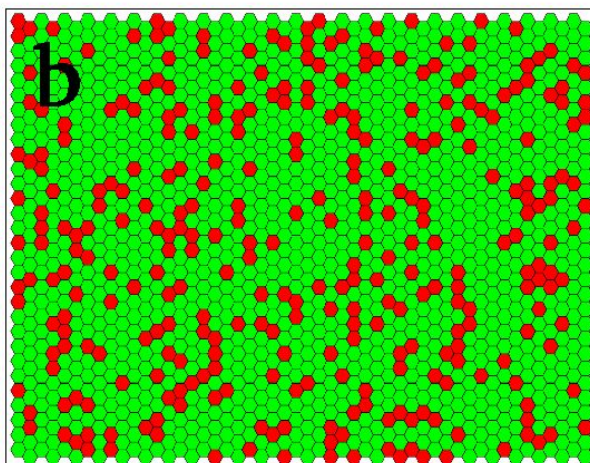
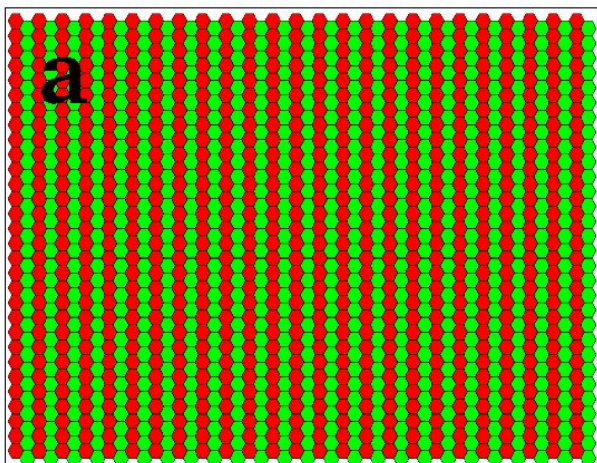


Figure 2:

Simulation results for a model fly eye of size 30X50 obtained via the general response function of our model (Eq. 6) with

- (a) $\alpha = 1$. This models a perfect Doli fly, where there are no mistakes.
- (b) $\alpha = 0$. This models Drosophila, where the two colours are randomly distributed, with a bias towards the green
- (c) $\alpha = 1$. This models a Doli fly with a perturbed column (eighth from right). Notice that the perturbations propagate leftward to the end, giving rise to a domain with a discernible boundary.
- (d) $\alpha = 0.7$. The parameters ε (local speckle correlation coefficient) = 0.95, $P_0^{Doli} = 0.0001$ and $P_0^{Droso} = 0.75$ were chosen with a view to match, at least qualitatively, the patterning of the intermediate, partially organised *Chrysosoma* sp. shown in Fig. 1d.

In fact, Eqs. 1 and 2 have a complex phase diagram as a function of their parameters, which is key to the richness of observable patterns in our model. Figure 3 is an illustration of the quantitative behaviour of these equations as a function of the parameters β , γ , P_0 and X_0 , with γ and X_0 scaled by β for convenience. The region of negative β corresponds to the uniformly patterned case, as mentioned above. For positive β , we fix the scaled threshold X_0 / β to be 0.2 for purposes of illustration, and plot the observed behaviours as a function of the remaining parameters. The blue and purple regions correspond to uniform retinal colours, while the green region is characterised by alternating-coloured stripes as found in Doli fly eyes. The white region is one where the biological factor X takes on negative values, which is physically inadmissible. In the upper right of the figure (*i.e.*, beyond $P_0 > 0.5$), the value of X_j will be either $\gamma - \beta(1 - P_0)$ or $\gamma - \beta P_0$ depending on its initial value. It is important to emphasise that the choice of threshold X_0 determines the specific numerical values at which some of these crossovers occur, so that the full phase behaviour is much richer than our simple representation in Figure 3.

Inter-column correlations and propagation of mistakes.

In the previous section, we dealt with the default probabilities of a column being of a particular colour, e.g. green; specifically, for a Doli retina, for a fully green column, this probability is 1 while it is 0 for a fully red column. However, if the column is not homogeneous, so that a few ommatidia (referred to henceforth as 'mistakes') are coloured differently from the majority, we need to add another ingredient to our model. This involves the influence of correlations: A green ommatidium in an otherwise fully red column will influence ommatidia that are yet to be formed in its neighbourhood to also be 'mistakes', i.e. to be red within the adjacent, predominantly green column. This neighbourhood will be referred to below as the correlation neighbourhood.

In quantitative terms, the probability is modified by a correction term l_{ij} due to the effect of mistakes, if any. Note that there are two indices now, since we are discussing the effect of corrections at the i^{th} ommatidium in the j^{th} column. First we define columnar types: a G column is one where most ommatidia are of the 1 type (green), while an R column is where most ommatidia are of the 0 type (red). A mistake is therefore defined as an element that takes a value 1 in an R column or 0 in a G column. Next, we define the correlation neighbourhood quantitatively: we postulate that correlations decrease exponentially with distance from the target site, so that only the 'nearest' mistakes in the previous column have a significant effect on the ommatidium under consideration. When this correction l_{ij} (see Materials and Methods, Appendix I for details) is added to the default probability, we have the expression for the full probability of the ommatidium ij being green:

$$P_{ij}(S,l) = p_j(S|X) + l_{ij} \quad (3)$$

The addition of the correlation factor thus brings about a (positive or negative) correction to the default value of the probability, due to the presence of mistakes in the adjoining column. If there are no mistakes in the earlier column, the default value of the probability is maintained. Figure 2c shows a simulation of a Doli eye with mistakes. Some of these

are isolated and have little long-distance effect (see picture of real Doli eye in Fig. 1b); on the other hand, there are occasions when, possibly due to an accident or changes in the thermal environment during development, there are arrays of mistakes extending over several sites in a column. In such a severely perturbed column, mistakes can then be propagated to adjoining columns much further away (Figure 1c).

We can model the effect of such a perturbed column in our formalism by setting $p_j = P_{\text{per}} = 0.5$, for a particular column j . The resultant simulation is shown in Figure 2c where the perturbed column is eighth from the right, and has clusters of red 'mistakes' in a column that is overall green. These mistakes are in close proximity to each other, leading to a propagation of errors for a considerable distance as the furrow moves onward (leftward, as the furrow propagates from right to left), sometimes never recovering before the end of the migration. We compare this simulation with a picture of a real Doli eye (Fig. 1c) where similar domains of mistakes form to the left of severely perturbed columns. By contrast, an isolated mistake (see centre of Fig. 1b) does not have long-term effects.

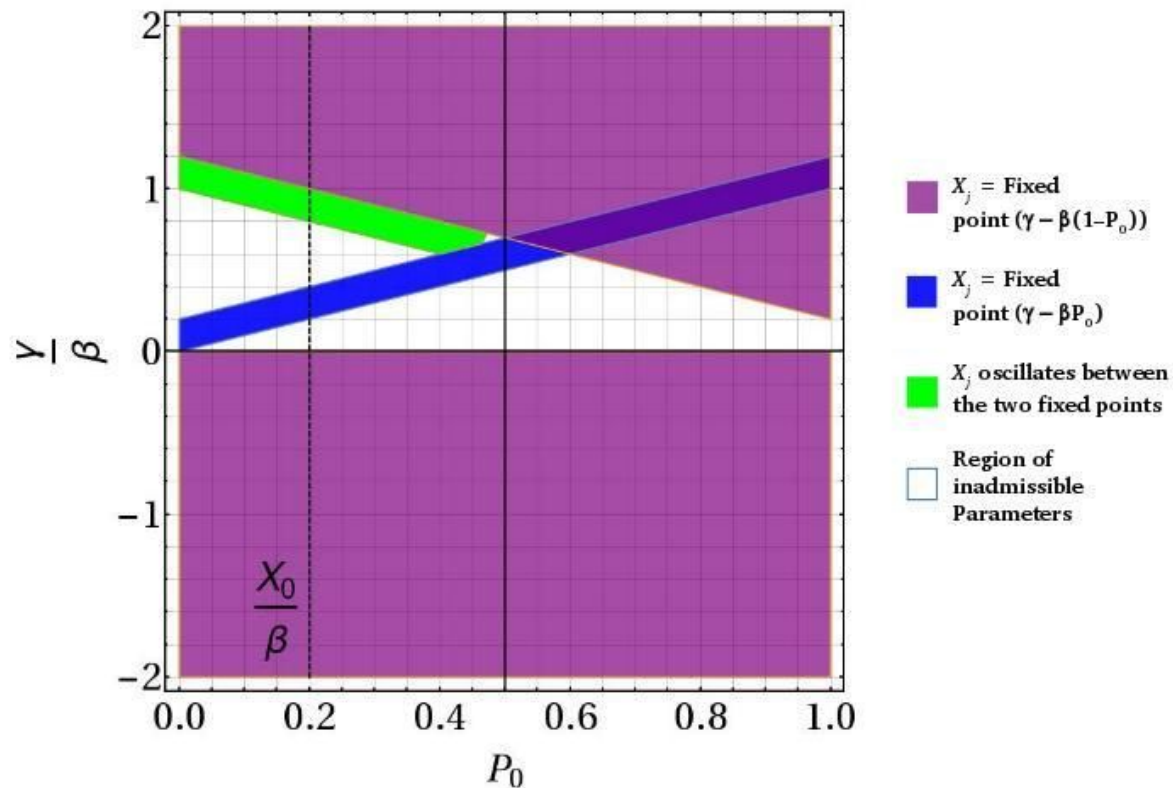


Figure 3: Sample plot of the phase behaviour of Eq. 2. The limit behaviour of X_j for all columns j , is plotted as a function of the parameters β , γ , P_0 and X_0 , with γ and X_0 scaled by β for convenience; the (scaled) threshold value of X_0 is set to a sample value of 0.2, as shown by the dotted line. For this choice, X_j approaches fixed point values $X_j = \gamma - \beta(1 - P_0)$ and $X_j = \gamma - \beta P_0$ respectively for the purple and blue regions. For these regions, the pattern is thus uniform, In the green region, X_j alternates between these two fixed points, giving rise to the stripes typical of Doli-like behaviour. The white region is physically inadmissible as X_j takes negative values (see text). In the top right where colours overlap, the limit values of X_j depend on their initial values; however, there is no alternation here, so that the behaviour is characteristic of the uniform fly.

Stochasticity in the Drosophila eye.

At the other end of the spectrum from Doli, the Drosophila retina contains ommatidia where R8 cells express either green or blue Rhodopsins that are randomly distributed, with a bias for green (65%) vs. blue (35%). This pattern is fully random (5, 22–24). This implies that the choice in every ommatidium is fully independent of all its neighbours, indicating that there is a total absence of correlations and full stochasticity: no auxiliary equation like Eq. 2 is therefore needed. On the other hand, the bias can be accommodated by choosing the default probabilities of the two subtypes to be unequal. The analogue of Eq. 3 in this case becomes:

$$P_{ij}^{Droso}(S, l) = p_j^{Droso}(S|X) \quad (4)$$

where P_{ij} is the probability that the ommatidium ij is green. Figure 2b shows a sample configuration simulated with $P_0 = 0.35$ (see Eq. 1) to be compared with an image of a Drosophila eye (Figure 1b).

Predicting the patterns of other fly eyes.

Our simple theory is able to come up with predictions for retinal patterns in fly species that have not yet been analysed. The simplest way of doing this is by mixing the characteristics of order and disorder while defining the colour probabilities of ommatidia. Specifically, we use a linear combination of random *Drosophila*-like and ordered uniform of *Doli*-like ommatidial colour probabilities with a coupling coefficient $\alpha \in [0,1]$:

$$P_j^\alpha(S|X) = \alpha p_j^{\text{OrderedFly}}(S|X) + (1 - \alpha) p_j^{\text{Droso}}(S|X), \quad (5)$$

$$P_j^\alpha(S|X) = \alpha p_j^{\text{Doli}}(S|X) + (1 - \alpha) p_j^{\text{Droso}}(S|X). \quad (6)$$

Eq. 5 applies for $\beta < 0$, and Eq. 6 for $\beta > 0$. As a test of this formalism, we try to model the eye of the *Chrysosoma* sp fly. As in the case of *Doli* eyes, the pattern is that of corneal colours rather than Rhodopsin expressed in photoreceptors, which are not known for this species. In the absence of detailed experimental studies of the geometrical correlations of these corneal colours, we restrict ourselves to generating a qualitatively similar pattern to that found in the fly (Figure 1d). Given the uniform (rather than striped background), we work with the negative β situation embodied by Eq. 5; our choice of a finite α leads to the blend of randomness and order shown in Figure 2d. This heuristic choice of α gives reasonable agreement with patterning in the real *Chrysosoma* sp fly; however, as we shall see below, our formalism allows for fully quantitative comparisons and insights where experimentally measured correlations of fly eye colours are available.

Horizontal and vertical correlation coefficients can be defined for our model as shown in Appendix II, Materials and Methods. A generic phase diagram (Figure 4) for fly eyes can thus be generated, which provides a powerful tool for analysis as follows: First, the phase diagram is predictive for fly eyes that may not yet have been found or analysed.

Next, detailed quantitative colour correlations computed for a particular fly can be placed at a particular point on the phase diagram, whose location would enable one to identify the genesis of the fly, i.e. whether it is derived from a perturbation of an ordered fly such as Doli, or a disordered one such as *Drosophila*, etc. When coupled with the theory presented above, this would then be a very powerful tool for the identification of specific mechanisms associated with the morphogenesis of an arbitrary, unknown fly type.

choice of α gives reasonable agreement with patterning in the real Soldier fly; however, as we shall see below, our formalism allows for fully quantitative comparisons and insights where experimentally measured correlations of fly eye colours are available.

Horizontal and vertical correlation coefficients can be defined for our model as shown in Appendix II, Materials and Methods. A generic phase diagram (Figure 4) for fly eyes can thus be generated, which provides a powerful tool for analysis as follows. First, the phase diagram is predictive for fly eyes that may not yet have been found or analysed. Next, detailed quantitative colour correlations computed for a particular fly can be placed at a particular point on the phase diagram, whose location would enable one to identify the genesis of the fly, i.e. whether it is derived from a perturbation of an ordered fly such as Doli, or a disordered one such as *Drosophila*, etc. When coupled with the theory presented above, this would then be a very powerful tool for the identification of specific mechanisms associated with the morphogenesis of an arbitrary, unknown fly type.

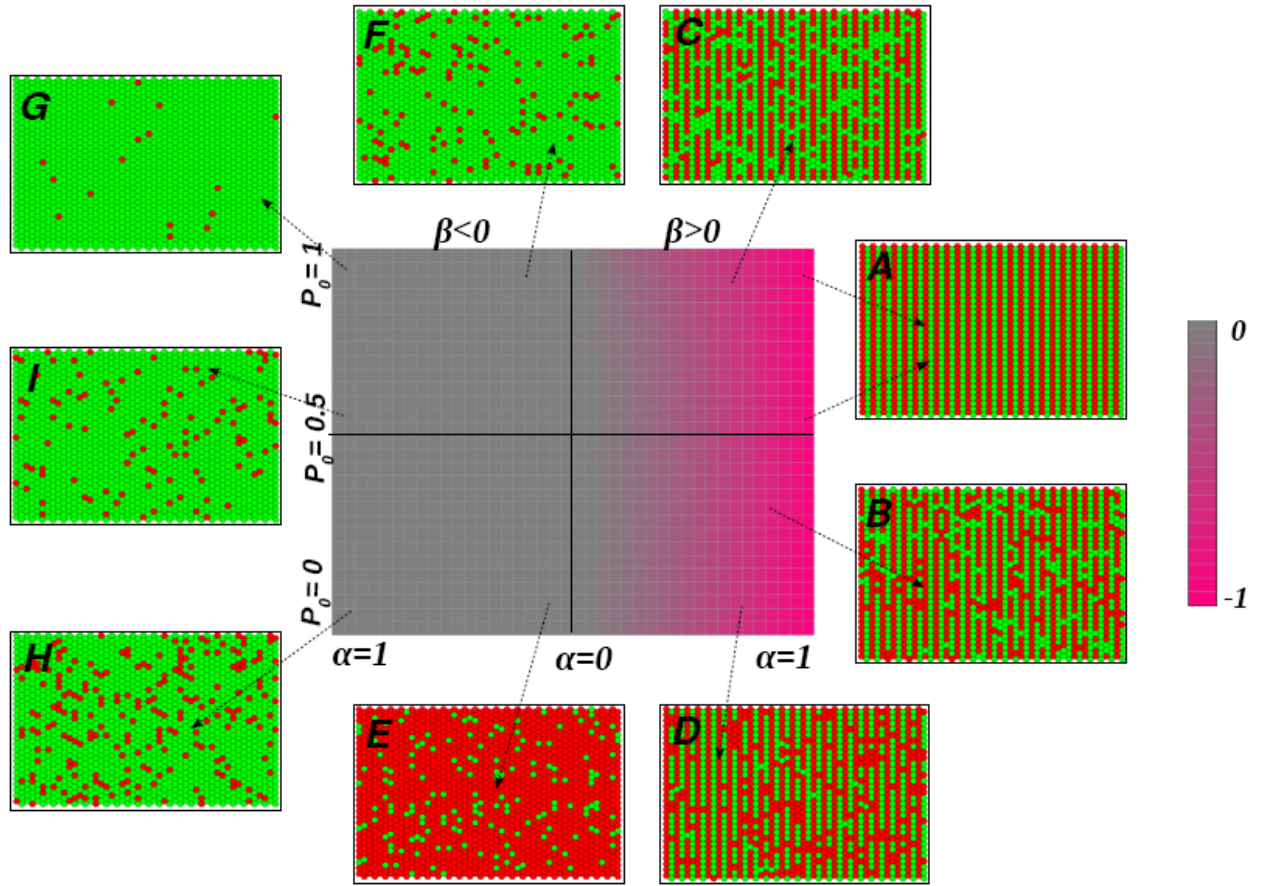


Figure 4: The mean horizontal correlation coefficient $\langle R^h \rangle$ diagram for general values of parameters (Eq. 15). The left half of the main diagram corresponds to $\beta < 0$ (Eq. 5) and the right half corresponds to $\beta > 0$ (Eq. 6). Along the x-axis, α varies from 0 to 1 to the right and the left of the origin; P_0^{Droso} is plotted along the y-axis. $\langle R^h \rangle = -1$ is denoted by pink, while $\langle R^h \rangle = 0$ is denoted by grey. The panels (A)-(I) show simulations of eyes corresponding to the parameter values to which they are connected by arrows as follows:

(A) $\beta > 0$, $\alpha = 1$. The value $\alpha = 1$ makes the second term in Eq. 6 vanish, so that the result is independent of P_0^{Droso} . The pure Doli pattern shown is thus obtained at both the extreme right (and the extreme left) of the diagram

(B) $\beta > 0$, $\alpha = 0.9$ and $P_0^{Droso} = 0.4$. The randomness of Drosophila permeates to some extent the order of Doli, and the absolute value of $\langle R^h \rangle$ decreases).

(C) $\beta > 0$, $\alpha = 0.65$ and $P_0 = 0.9$. As α decreases, mistakes become more frequent, although the stripes of Doli are still distinguishable.

(D) $\beta > 0$, $\alpha = 0.65$ and $P_0 = 0.1$. This is the complement of C, with the difference that there

is a higher probability of red. **(E)** and **(F)** have values of α close to zero, so that a *Drosophila* dominant behaviour is observed with $\langle R^h \rangle = 0$. Again these are complementary, $P_0^{Droso} = 0.1$ and 0.9 resulting in a small and large percentage of green respectively (G), (H) and (I): $\alpha = 0.9$, with $P_0^{Droso} = 0.9$ **(G)**, $P_0^{Droso} = 0.55$ **(H)** and $P_0^{Droso} = 0.1$ **(I)**. The green (red) colour or the ordered fly is dominant (90%) and the value of P_0^{Droso} determines the colour of the 10% ($1 - \alpha$) impurities.

Discussion

Statistical physics can be usefully applied to biology when searching for organizational principles (25–27). The present study is such an example: we present a minimal model that describes patterns observed in the retinas of *Drosophila* and *Dolichopodidae* flies, and via a generic phase diagram, we are able to predict parameters that might underlie a range of patterns found on the eyes of various fly species. The essential ingredients of this are stochasticity and correlations: an ommatidium chooses its colour state depending on the competition between its gene-expression-induced probability to be a specific colour (e.g. the expression of *Spineless*) and the influence of its prior column along the progression path of the morphogenetic furrow. Intercolumnar correlations greatly outweigh the effect of stochasticity in *Doli*, while the reverse is the case for *Drosophila*. This formalism also allows us to probe how mistakes propagate in the retina, for example in *Doli*. For instance, supposing an ommatidium in a column is a ‘mistake’, i.e. is of the wrong colour, our formalism suggests that two scenarios are possible. If the mistake is isolated, it will soon be resolved as the morphogenetic furrow progresses. However, a cluster of mistakes can lead to interesting domain formation, with interfacial roughness as in, say, crystalline systems where dislocations define the area of interfacial disorder (28). This is in accord with experimental observations (5, 20).

The richness of our formalism, however, goes well beyond the description of known species to the prediction of those that are as yet undiscovered. In terms of simple model parameters, we are able to generate a global phase diagram for flies with binary colour choices and allow for important clues to the morphogenetic mechanisms at play. The suggested parameter values for the *Chrysosoma* species are an example of this power, although we emphasise that in the absence of quantitative experimental data about the molecules involved and variables such as their diffusion constants, we have here aimed at only qualitative agreement. In fact, we hope this work will motivate detailed quantitative analysis of experimentally observed patterns, as well as genetic analysis of factors involved in the expression of colour in fly eyes, so that our predictions can be put to the test. Once our model is adequately tested for two-colour retinæ, we will be able to extend our analysis to flies and other insects with more complex colour patterning.

Supplementary Data – Materials and Methods

Appendix I -- Calculation of coupling l_{ij} for error propagation.

The propagation of errors is described by an additive probability term l_{ij} for each site, (cf. Equation 3) in the main text. This term is defined as

$$l_{i,j+1} = \varepsilon \sum_{h \in N((i,j)/2)} [a_{h,j} - \Theta(p_j - p_{j+1})] \exp\left(-\frac{(i-h)^2}{k}\right) \quad (7)$$

with the step function

$$\Theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Assuming a stripe pattern, the Θ -term in Eq. (7) gives the a -value expected to be dominant in column j . For $p_j > p_{j+1}$, the probability of green ommatidia ($a_{i,j} = 1$) in column j is larger than that in the adjacent column j . Green ommatidia are the standard in this column j , and a red ommatidium $a_{i,j} = 0$ in a row i is considered an error. Likewise, green ommatidia are considered errors in red-dominated columns ($p_j < p_{j+1}$), where preceding the description holds analogously by exchanging 0 and 1 entries of a . An error in column j potentially perturbs the nearby ommatidia in the following column $j+1$. The influence is assumed to decay exponentially with the distance between ommatidia; hence the exponential factor with a length scale parameter k in Eq. (7). Influences from all errors in the previous column sum up linearly (summation with row index h) and are scaled with a coupling strength ε . Accounting for the hexagonal lattice, we distinguish two index sets for the row index. The set

$$N(0) = \{1, 2, \dots, n\} \quad (9)$$

is the usual set of integer index values, used for the columns j with j even. In columns with odd j , sites have row indices in

$$N(1) = \{1.5, 2.5, 3.5, \dots, n - 0.5\}. \quad (10)$$

Appendix II -- Analysis of correlations in patterns.

Given a pattern (a_{ij}) with row index $i \in \{1, \dots, n\}$ and column index $j \in \{1, \dots, m\}$, the horizontal autocorrelation coefficient is defined as

$$R^h = \sigma^{-2} n^{-1} (m-1)^{-1} \sum_{i=1}^n \sum_{j=1}^{m-1} (a_{ij} - \mu)(a_{i,(j+1)} - \mu) \quad (11)$$

with the pattern's mean value

$$\mu = (nm)^{-1} \sum_{i=1}^n \sum_{j=1}^m a_{ij} \quad (12)$$

and variance

$$\sigma = -\mu^2 + (nm)^{-1} \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2. \quad (13)$$

Likewise, the vertical autocorrelation coefficient is defined as

$$R^v = \sigma^{-2} (n-1)^{-1} m^{-1} \sum_{i=1}^{n-1} \sum_{j=1}^m (a_{ij} - \mu)(a_{(i+1),j} - \mu). \quad (14)$$

Since we are studying stochastic pattern generation, r realizations ($r \gg 1$) are performed for a given set of parameter values (α, β, P_0) . For the pattern generated in each realization k ($1 \leq k \leq r$), the autocorrelation coefficients R_k^h and R_k^v are

computed. Then these coefficients are averaged

$$\langle R^h \rangle = r^{-1} \sum_{k=1}^r R_k^h, \quad \langle R^v \rangle = r^{-1} \sum_{k=1}^r R_k^v. \quad (15)$$

References

1. Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135(2):216–26.
2. Little SC, Tikhonov M, Gregor T (2013) Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. *Cell* 154(4):789–800.
3. Boettiger AN, Levine M (2009) Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science* 325(5939):471–3.
4. Lagha M, et al. (2013) Paused Pol II coordinates tissue morphogenesis in the *Drosophila* embryo. *Cell* 153(5):976–987.
5. Wernet MF, et al. (2006) Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature* 440(7081):174–80.
6. Koch AJ, Meinhardt H (1994) Biological pattern formation: from basic mechanisms to complex structures. *Rev Mod Phys* 66(4):1481–1507.
7. Hardie RC (1985) Functional Organization of the Fly Retina. *Progress in Sensory Physiology*, Progress in Sensory Physiology., eds Autrum H, et al. (Springer Berlin Heidelberg, Berlin, Heidelberg). doi:10.1007/978-3-642-70408-6.
8. Rister J, Desplan C (2011) The retinal mosaics of opsin expression in invertebrates and vertebrates. *Dev Neurobiol* 71(12):1212–26.
9. Franceschini N, Kirschfeld K, Minke B (1981) Fluorescence of photoreceptor cells observed in vivo. *Science* 213(4513):1264–7.
10. Johnston RJ, Desplan C (2010) Stochastic mechanisms of cell fate specification that yield random or robust outcomes. *Annu Rev Cell Dev Biol* 26:689–719.
11. Perry M, et al. (2016) Molecular logic behind the three-way stochastic choices that expand butterfly colour vision. *Nature* 535(7611):280–284.
12. Wernet MF, Perry MW, Desplan C (2015) The evolutionary diversity of insect retinal mosaics: common design principles and emerging molecular logic. *Trends Genet* 31(6):316–328.
13. Roignant J-Y, Treisman JE (2009) Pattern formation in the *Drosophila* eye disc. *Int J Dev Biol* 53(5–6):795–804.
14. Yamaguchi S, Wolf R, Desplan C, Heisenberg M (2008) Motion vision is independent of color in *Drosophila*. *Proc Natl Acad Sci U S A* 105(12):4910–5.
15. Yamaguchi S, Desplan C, Heisenberg M (2010) Contribution of photoreceptor subtypes to spectral wavelength preference in *Drosophila*. *Proc Natl Acad Sci U S A* 107(12):5634–9.
16. Wernet MFF, et al. (2011) Genetic Dissection Reveals Two Separate Retinal Substrates for Polarization Vision in *Drosophila*. *Curr Biol* 22(1):12–20.
17. Lubensky DK, Pennington MW, Shraiman BI, Baker NE (2011) A dynamical model of ommatidial crystal formation. *Proc Natl Acad Sci U S A* 108(27):11145–50.

18. Huber A, et al. (1997) Molecular cloning of *Drosophila* Rh6 rhodopsin: the visual pigment of a subset of R8 photoreceptor cells ¹. *FEBS Lett* 406(1–2):6–10.
19. Wernet MF, et al. (2003) Homothorax switches function of *Drosophila* photoreceptors from color to polarized light sensors. *Cell* 115(3):267–79.
20. Jukam D, Lidder P, Desplan C (2008) Rhodopsins in *Drosophila* Color Vision. *Visual Transduction and Non-Visual Light Perception* (Humana Press, Totowa, NJ), pp 251–266.
21. Johnston Jr. RJ, Desplan C, Johnston RJ, Desplan C (2014) Interchromosomal communication coordinates intrinsically stochastic expression between alleles. *Science* (80-) 343(6171):661–665.
22. Johnston RJ, et al. (2011) Interlocked feedforward loops control cell-type-specific Rhodopsin expression in the *Drosophila* eye. *Cell* 145(6):956–968.
23. Bell ML, Earl JB, Britt SG (2007) Two types of *Drosophila* R7 photoreceptor cells are arranged randomly: A model for stochastic cell-fate determination. *J Comp Neurol* 502(1):75–85.
24. Wolfram S (1984) Cellular automata as models of complexity. *Nature* 311(5985):419–424.
25. Rohlf T, Bornholdt S (2009) Morphogenesis by coupled regulatory networks: reliable control of positional information and proportion regulation. *J Theor Biol* 261(2):176–93.
26. Wolfram S (1983) Statistical mechanics of cellular automata. *Rev Mod Phys* 55(3):601–644.
27. Snyder AW (1979) Physics of Vision in Compound Eyes. *Physics of Vision in Compound Eyes* (Springer Berlin / Heidelberg), pp 225–313.
28. Shinde DP, Mehta A, Barker GC (2014) Shaking-induced crystallization of dense sphere packings. *Phys Rev E* 89(2):22204.